

## Electron Capture/Transfer versus Collisionally Activated/Induced Dissociations: Solo or Duet?

Roman A. Zubarev, Alexander R. Zubarev, and Mikhail M. Savitski

Division of Molecular Biometry, Department of Cell and Molecular Biology, Uppsala University, Uppsala, Sweden

New ion fragmentation technologies—electron capture dissociation (ECD) and electron-transfer dissociation (ETD)—are based on interaction of multiply charged polypeptides with either free electrons (ECD) or anionic species (ETD). After initial difficulties, these ECD/ETD (ExD) technologies are now being increasingly implemented in high-throughput proteomics work. This critical analysis presents arguments for the combined use of ExD with the conventional low-energy collisional excitation CID/CAD (CxD). It is argued that the database search, a key technology in MS/MS-based proteomics, is vulnerable with respect to the incomplete sequence information obtainable with either of the techniques, peptide MS/MS homology being a major complicating factor. De novo sequencing is viewed as the only adequate answer to this challenge and it can be achieved only with combined use of ExD and CxD. The payoff in the form of additional sequence information is projected to exceed the costs of such implementation. The greatest impact of combining ExD and CxD is expected in high-resolution instruments. (J Am Soc Mass Spectrom 2008, 19, 753–761) © 2008 American Society for Mass Spectrometry

For several years after the emergence of electron capture dissociation (ECD) a decade ago [1] it has not been clear whether this technology could be used in high-throughput analysis. Originally, ECD has been available only in Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometry, where the typical reaction time with electrons was 10 s and integration of at least 10 acquisitions was needed for achieving acceptable signal-to-noise ratios [2]. Switching from filament-based electron sources to indirectly heated dispenser cathodes reduced the irradiation time to milliseconds and improved the fragment collection efficiency [3, 4]. However, the issue of low signal-to-noise ratios remained and, since a single spectrum acquisition in FT-ICR lasts several hundred milliseconds or longer, integration of several scans was a severe throughput limitation. Despite a few reports on successful on-line implementation [5–7], ECD remained essentially an off-line technique. Later, a combination of optimized FT-ICR cell design, on-axis filament arrangement and low-energy gated trapping allowed one to achieve good-quality ECD mass spectra in a single acquisition [4], which opened a way for high-throughput analysis. Several groups worked on implementation of ECD in radiofrequency trapping devices, with encouraging first

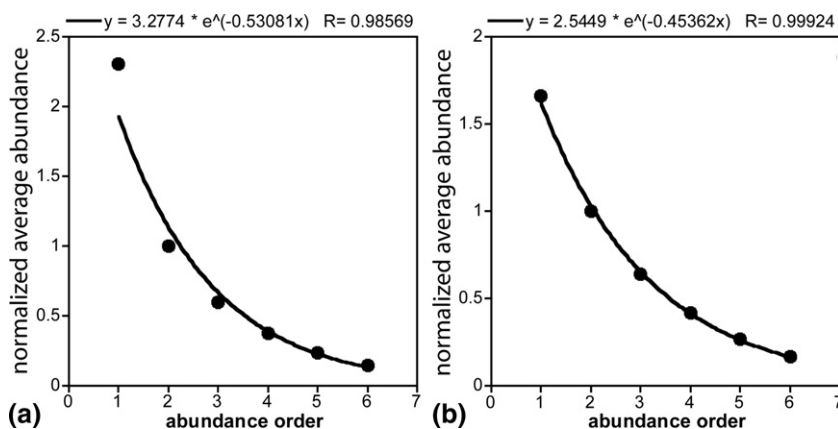
results [8, 9]. However, the real breakthrough in making ECD-type technology available to a broader circle of researchers came with the discovery of efficient ion–ion reactions with radical anions termed electron-transfer dissociation (ETD) [10]. The evidence of that is the string of recent high-profile publications [11–14]. In view of this success, the role of ECD/ETD (which we shorten to ExD) in mass spectrometry deserves a separate discussion. To initiate such a discussion, we chose as a central topic the relationship between the traditional collisional excitation techniques collisionally activated dissociation (CAD)/collision-induced dissociation (CID) (CxD) and ExD in the most frequent in the proteomics task of peptide sequence analysis. Is there a conflict, competition, or cooperation between these two groups of fragmentation techniques? Will they evolve hand by hand or will one of them become dominant and the other one doomed to extinction? Below we try to address these issues, in no way pretending to cover all relevant aspects.

### Discussion

#### *Is CAD/CID Alone Sufficient?*

If CxD could solve all analytical problems encountered in mass spectrometry, no additional/alternative fragmentation technique would be needed. The reasons that CxD is insufficient in proteomics have recently been

Address reprint requests to Prof. Roman A. Zubarev, Uppsala University, Division of Molecular Biometry, Institute for Cell and Molecular Biology, Laboratory for Biological and Medical Mass Spectrometry, Box 598, SE-751 24 Uppsala, Sweden. E-mail: [roman.zubarev@icm.uu.se](mailto:roman.zubarev@icm.uu.se)



**Figure 1.** (a) distribution of normalized average  $y$ -ion abundances in >15,000 CAD mass spectra of doubly protonated tryptic peptides from the SwedCAD database [17]. The curve is the best exponential fit. (b) Same as (a) except that all peptides containing proline are discarded, and  $y_{n-2}$  fragments are not considered.

discussed [15]. Briefly, there are two groups of reasons. One is that CxD does not cleave nearly all the required bonds to derive full information on primary structure of a polypeptide [16]. The other is that, even when all peptide bonds are cleaved, the obtained information on the fragment masses is still insufficient to uniquely identify the sequence and its modifications. Let us investigate these reasons more closely.

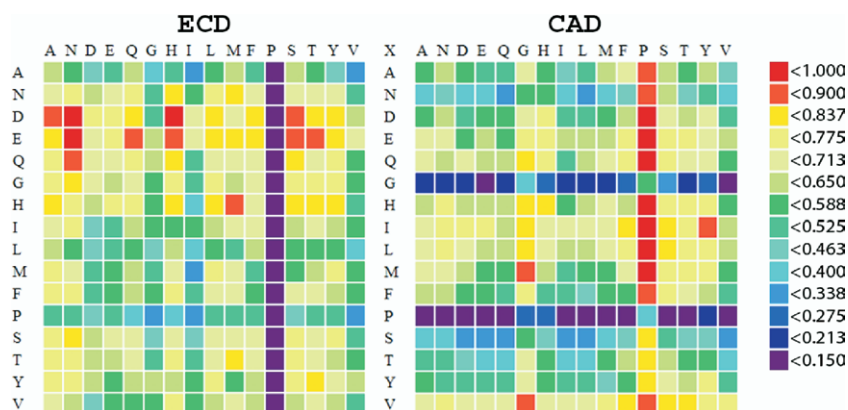
To understand why CxD does not in the general case cleave all peptide bonds, consider the distribution of CxD peptide  $y$ -ion fragment abundances (Figure 1a). These abundances are obtained in fragmentation in a linear ion trap of doubly charged tryptic peptides [17]—that is, where both the size of the molecules (about 10 residues on average) and the ratio between the number of ionizing protons (two) and strongly basic groups (one) is advantageous for obtaining abundant backbone fragmentation [18, 19]. Even in that almost ideal situation, however, the fragment abundances rapidly fall off, with the fifth (Figure 1a) or sixth (Figure 1b) most abundant fragment being on average 10 times less intense than the fragment with the highest abundance. There is a slight anomaly in the ratio between the first two peaks in Figure 1a, which is eliminated (Figure 1b) when proline-containing peptides are disregarded, along with the special case of  $y_{n-2}$  cleavage. A 13-residue peptide requires 12 backbone cleavages for complete primary structure determination; it can be deduced that, if the trend in Figure 1 continues, the least abundant of these cleavages will on average be 100 times less abundant than the most abundant fragment. Therefore, if the signal-to-noise (S/N) ratio of the precursor ion upon isolation before MS/MS is less than 200:1, one cannot in general expect a complete cleavage pattern (assuming S/N 2:1 as a threshold for fragment detection). The latter estimate is done based on 100% conversion efficiency of the precursor ions into backbone fragments, which is an idealization (only half of the total ion abundance in MS/MS spectra obtained in an

ion trap can be attributed to  $b$ - or  $y$ -ions [18]. Taking into account less than 100% fragment detection efficiency and the fact that singly charged fragments produce in induced-current detectors (the type used in Fourier transform mass spectrometry) half as strong a signal as the same number of doubly charged precursors, we conclude that S/N 1000:1 is a realistic value of the lowest signal required for complete sequencing by CxD of moderately large doubly charged tryptic peptides.

The requirement of the high signal abundance is a significant analytical limitation of CxD. The cause of this limitation is the “mobile proton” mechanism of CxD backbone cleavages [19] that is dominant in multiply protonated tryptic peptides with one strongly basic group. The simplest model quantitatively describing the “mobile proton” fragmentation [20] suggests that the cleavage probability *after* an amino acid residue is determined by the probability for a proton to reside on the respective carbonyl, which in turn is determined by the carbonyl basicity GB of this residue. Within that model, the “dynamic range” of CxD cleavages is determined by the difference  $\Delta GB_{\max}$  between the maximum and minimum GB values divided by  $kT_{\text{int}}$ , where  $T_{\text{int}}$  is the internal temperature. Thus higher  $T_{\text{int}}$  should narrow the dynamic range. Such an increase can be achieved by more energetic ion-neutral collisions. However, higher collision energies usually lead to a larger scattering and often result in a decrease in the efficiency of fragment collection and thus in an overall loss of sensitivity.

In the absence of mobile protons, different CxD mechanisms prevail [21], in which carbonyl basicities play a less important role compared to the chemical activity of side chains. The “dynamic range” of such cleavages can be even greater than that in the presence of a mobile proton [18].

Limited CxD efficiency in sequencing of polypeptides is also caused by a frequent overlap of the masses of N-terminal and C-terminal fragments. Such an over-



**Figure 2.** Amino acid preferences in >15,000 tandem mass spectra CAD and ECD obtained for the same peptides as in Figure 1. Compared to Savitski et al. [23], the database was extended, data for Met were added.

lap can lead to mistakes in spectrum interpretation [22]. The overlap does not have to be precise; even partial fusion of isotopic distributions can confuse computer algorithms performing peak assignment. Higher resolution and mass accuracy always reduce the risk of an error, but are only a partial cure since theoretical masses of *b*- and *y*-ions can be as close as desired and even identical. Thus the risk of an overlap is always present and it grows with the molecular mass.

Another important factor influencing CxD efficiency is that low-energy collisions do not allow one to distinguish between the isomeric Leu/Ile (Xle) residues. Together these residues account for 1/6 of all amino acids found in natural proteins, which means that any peptide more than four residues long more likely than not contains at least one Xle residue. The amount of Xle isoforms (groups of tryptic peptides with identical sequences except for Leu/Ile permutation) in protein databases is very large; in our experience such isoforms are one of the main causes of wrong peptide assignment in a MS/MS database search. Note that neither improved CxD efficiency nor extremely high mass accuracy is capable of resolving Xle isoforms; only secondary, side-chain fragmentation resulting into *d*- and *w*-ions can help.

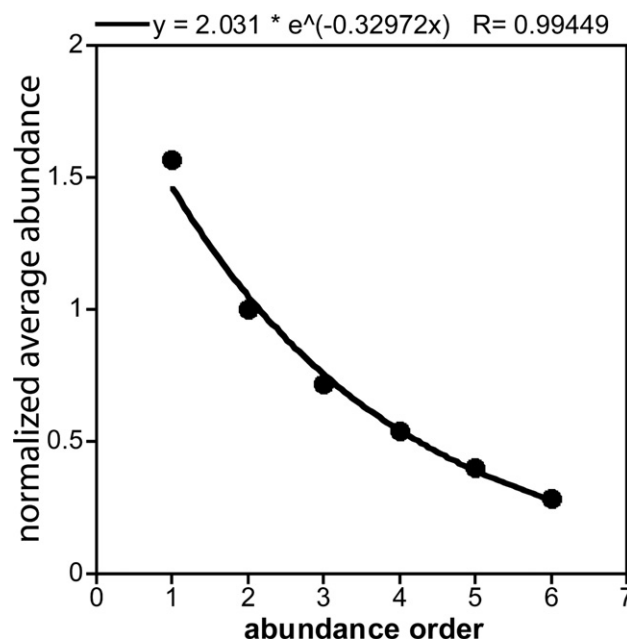
Finally, many posttranslational modifications in CxD are labile in the gas phase, such as Ser, Thr, and His phosphorylations as well as N- and particularly O-glycosylations. This makes CxD unsuitable for their characterization without chemical derivatization.

### Complementarity of ExD and CxD

In ExD, the N—C<sub>α</sub> bond is cleaved as opposed to the C—N bond in CxD, but this is not the main reason for the complementarity of these two techniques. To be truly complementary, ExD should preferentially cleave at different residues than CxD. Previously, statistical analysis has shown that this is indeed the case [23]; the conclusion is now confirmed with an updated and extended database. Figure 2 demonstrates these updated propensities

for CAD and ECD cleavages in a color-coded mode. There is no statistical correlation between the amino acid preferences in these two techniques (correlation coefficient is close to zero), and for the same peptide the site of the most abundant cleavage in CxD is usually shifted with respect to the similar site in ExD [23].

The “dynamic range” of ExD mass spectra is narrower than that of CxD (Figure 3). Since the internal temperatures in ExD are less than those in CxD, the narrower dynamic range means in a simple quantitative model a smaller  $\Delta G_{\max}$  value. The nature of the “G” values responsible for the cleavage frequencies in ExD is not yet understood. They are unlikely to relate to backbone amide basicities [20] and more likely to hydrogen atom affinities. Empirically, the best correlation



**Figure 3.** Distribution of normalized average *z*-ion abundances in >15,000 ECD mass spectra obtained for the same peptides as in Figure 1. The curve is the best exponential fit.

with ECD cleavage propensities was found for certain hydrophobicity indexes [24], but this correlation has so far found no simple explanation.

The advantage of ExD in narrow dynamic range of fragmentation compared to CxD increases its relative importance when molecular mass increases. Thus ExD is particularly suitable for fragmentation of larger polypeptides, including intact proteins [25–27].

The ability of ExD to differentiate between Ile and Leu residues via secondary fragmentation in radical  $z\bullet$  ions increases its utility as a complement to CxD [28–29]. Finally, the “soft” character of ExD cleavages makes this technique suitable for analysis of labile posttranslational modifications [30–32]. The complementarity of CxD and ExD is also reflected in the “small fragment” area. Whereas CxD produces immonium ions that are characteristic of the presence of certain amino acid in sequence, ExD gives few such ions but instead produces losses of neutral particle from the charge-reduced molecular species that are amino acid specific [33].

### Can ExD Replace CxD?

There is an obvious practical advantage in using just one fragmentation technique as a universal MS/MS tool, provided this tool satisfies analytical needs. Since it is unlikely that CxD (at least in its currently dominant form) would ever be able to replace ExD, let us consider the consequences of replacing CxD with ExD. The instrumental complexity and cost are likely to increase because CxD is easier to implement on most instruments than ExD. In shotgun proteomics, trypsin will no longer remain the most advantageous enzyme for peptide production because tryptic peptides tend to be preferentially doubly charged in electrospray ionization, which is a disadvantage for ExD where efficiency

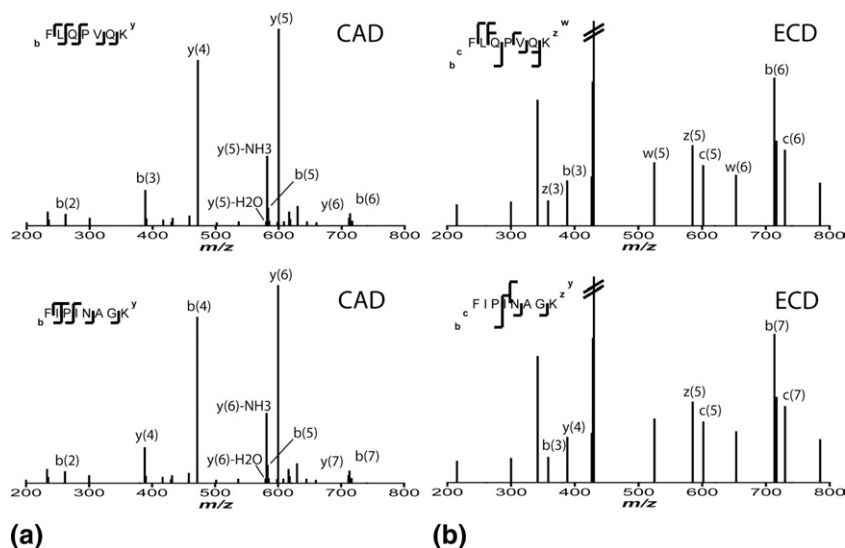
diminishes because of partial charge neutralization [34]. Of course, ExD is not applicable to singly charged ions or to negatively charge species (electron excitation dissociation, EED [35], and electron detachment dissociation, EDD [36], are special cases). Trypsin can of course be replaced by Lys-C or Arg-C [14, 37], although these enzymes tend to be less available. Importantly, Lys-C and Arg-C produce fewer peptides than trypsin, which reduces the sample complexity but may decrease the detection probability for low-abundance proteins. CxD of larger peptides produced by Arg-C and Lys-C is relatively less informative than CxD of smaller tryptic peptides, and thus the enzyme change will negatively affect the ability of this technique to complement ExD data. ExD seems to be sensitive to charge density in polypeptides, with efficiency dropping off with  $m/z$  faster than that in CxD [34]. Last but not least, most search engines are optimized for CxD MS/MS spectra and are less effective in processing ExD mass spectra.

One should also bear in mind that, whereas masses of  $b$  and  $y$  fragments in CxD are well-defined, masses of  $c$ - and  $z$ -ions in ExD can be 1 Da lighter or heavier, respectively, because of the extensive hydrogen rearrangement [2, 38, 39]. This effect reduces the certainty of fragment assignment from ExD-only mass spectra, which can be particularly challenging in de novo sequencing [16].

Thus we do not believe that the ExD-only approach is any more advantageous than the CxD-only regime. In our opinion, ExD and CxD technologies are best when implemented together, so that their complementarity can be exploited in the best possible manner.

### Combined Use of CxD and ExD

To appreciate the importance of complementary information, consider the CAD mass spectrum in Figure 4a.



**Figure 4.** MS/MS spectra of a doubly protonated peptide with fragments assigned to the sequence FLQPVQK (top) and FIPNAGK (bottom) obtained with: (a) CAD; (b) ECD.



In a Mascot search against the whole MSDB database using a  $\pm 20$  mDa mass accuracy window for fragment masses and  $\pm 3$  ppm for molecular masses, it received identical scores for both FLQPVQK and FIPINAGK peptides because both sequences are equally well explained by the MS/MS spectrum. Observe how some fragments are fitted as  $y$ -ions in one case and  $b$ -ions in the other case. However, the ECD spectrum (Figure 4b) shows a clear preference for the FLQPVQK sequence. Although technically the only difference in the backbone cleavage in ECD compared to CAD is the presence of the C-terminal cleavage to proline, this along with the two  $w$ -ions confirming the glutamine and leucine assignments [40] is enough to make a reliable distinction between the two sequences.

This example illustrates the nonlinear effect on uncertainty reduction of additional sequence information. Every additional piece of sequence information is of an increasing value, and the reverse is also true. For instance, in high-performance de novo sequencing, full sequences are far more reliable than sequences derived with “gaps,” even small gaps two residues long [16]. The nonlinear effect of complementary sequence information is worth remembering when assessing the cost/benefit balance of using two fragmentation techniques versus one.

### Is De Novo Sequencing Really Needed?

The benefits of complementary sequence information are particularly obvious in de novo sequencing [41]. However, there is a regrettably widespread perception that de novo sequencing of polypeptides is no longer required, given that a database search is very reliable and the volume of databases is rapidly expanding. In our opinion, the last two statements are in contradiction because the more limited the database is, the less information is required to identify a unique match in that database. However, finding a unique match in a limited database is not a guarantee of the validity of that match, and artificial reduction of the database size to “increase” the search validity is opposed by database search experts [42]. Below we argue that only de novo sequencing can guarantee error-free sequence identification.

### MS/MS Peptide Homology

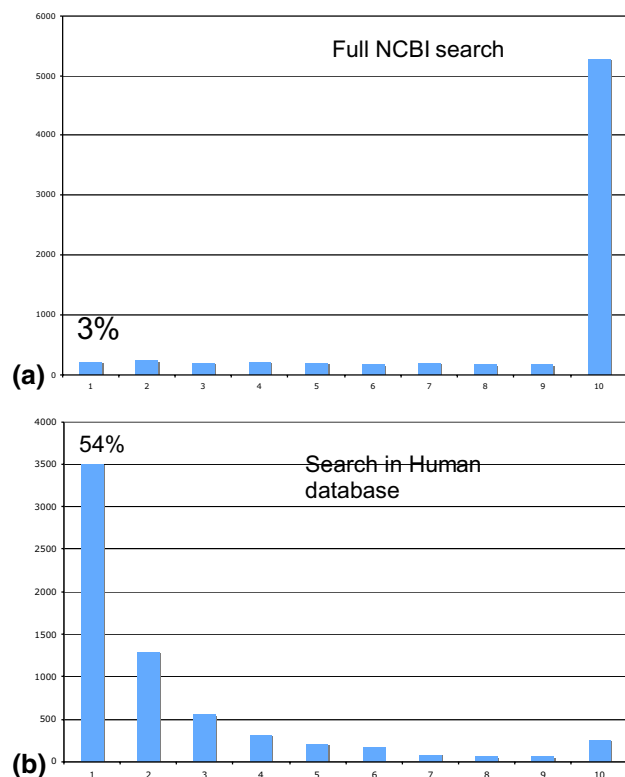
It is sometimes argued that Cx/D alone can identify peptide with any desirable certainty. There is no question that beautiful examples of such identification can be demonstrated. Proteomics is, however, a game of large numbers. When a significant amount of datasets just barely make it above the identification threshold (which is often the case), it is extremely difficult to provide the error-free status. Even 99.9% certainty in protein identification gives a few false positives when several thousand of proteins are identified. A reality

check is therefore advised when an exceptionally low false discovery rate is reported.

One of the serious limitations of the database search approach is peptide MS/MS homology, which in this discussion is not equivalent to sequence homology (the homology used in BLAST search), but relates to the fact that mass lists of MS/MS fragments of peptides with different sequences may look similar or even identical (see Figure 4a for an example). For instance, all peptides are MS/MS homologs of their own reversed sequences, given that  $b_n$  ions of the forward sequence coincide with  $(y_n' - H_2O)$  ions of the reversed sequence and  $y_n'$  ions of the forward sequence have the same masses as  $(b_n + H_2O)$  ions [43] of the reversed sequence. The larger the database, the larger is the chance of matching a homolog. The ultimate database includes all possible sequences and search in such a database is equivalent to de novo sequencing.

To illustrate the widespread appearance of peptide MS/MS homology, consider the results of a Mascot search of MS/MS data on peptides from the whole proteome of the human cell lines A431 performed in our lab some time ago [44]. Matrix Science warns against narrowing taxonomy in Mascot searches [42] and thus it may be prudent, as a first step of analysis, to perform a search in the full database for all species. Such a search may reveal the presence of, for example, *Mycoplasma* (between 5 and 35% of cell cultures in current use are probably infected with mycoplasmas [45]). The search in the full NCBI database gave a single, unique hit (with any score) for only 3% of MS/MS queries that gave any hit at all, and even narrowing the taxonomy to human-only gave more than one hit in 46% of the cases (Figure 5). Many MS/MS datasets gave repetitive, same-score top hits: 64% of queries in all-NCBI search and 20% in human-only search yielded *different* peptide sequences with the same top score (Figure 6).

Homological hits, in turn, come in two kinds. When an MS/MS spectrum receives a number of nonidentical sequence assignments with the same or very close score, no meaningful differentiation between the assignments can be made based on the score value only. In this case the search engine is usually asked to perform “clustering,” that is, to assign the MS/MS dataset to the protein already supported by the largest number of other MS/MS spectra. This is a sensible approach when a few abundant proteins are present in the sample. However, such an ideal situation is rarely met in real-life proteomics. In a more realistic case, there is a great multitude of low-abundance proteins in the sample. In such a situation, the “clustering” option often forces the search engine to make a difficult choice between assigning an MS/MS spectrum to a protein with no other detected peptides or to a protein with one low-scoring supporting peptide. The statistical validity of the “clustering” approach becomes in that case rather questionable. The clustering method may be responsible for the low overlap



**Figure 5.** Distribution of the number of peptide sequences (not necessarily unique) belonging to different proteins assigned by Mascot with any score to each MS/MS dataset: (a) for search in the full NCBI database; (b) for search against the human database. Ten was the highest allowed number of sequences.

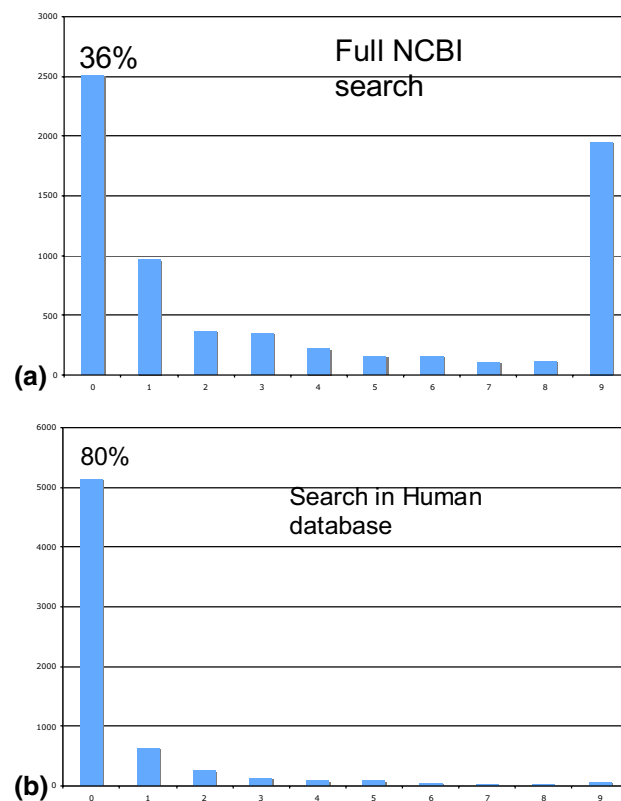
between the proteomes of the same biological system measured by different labs [46].

Another side effect of clustering is underestimation of the false discovery rate (FDR) when for the purpose of FDR estimation both “forward” and “reversed” databases are concatenated into a single database. For instance, matching the LC/MS/MS dataset from Nielsen et al. [44] against a reversed NSBI database gave 30 above-threshold hits. A subsequent search using a forward database revealed that 29 of 30 peptides also gave hits there. In 18 cases, the scores for the forward and reverse assignments were the same. In one of 18 cases, the sequence was palindromic, and thus both forward and reverse assignments were correct. The remaining 17 cases were expected to have statistically equal chances to be attributed to the forward and reverse databases in a concatenated search. However, in reality only two instead of the expected 8 to 9 peptides were attributed to the reversed database, with 15 assigned to forward sequences. Although the clustering function responsible for attributing some of the “extra” peptides to forward protein sequencing might have been right in some cases, it is unlikely that all of the forward assignments of same-score datasets were fully justified.

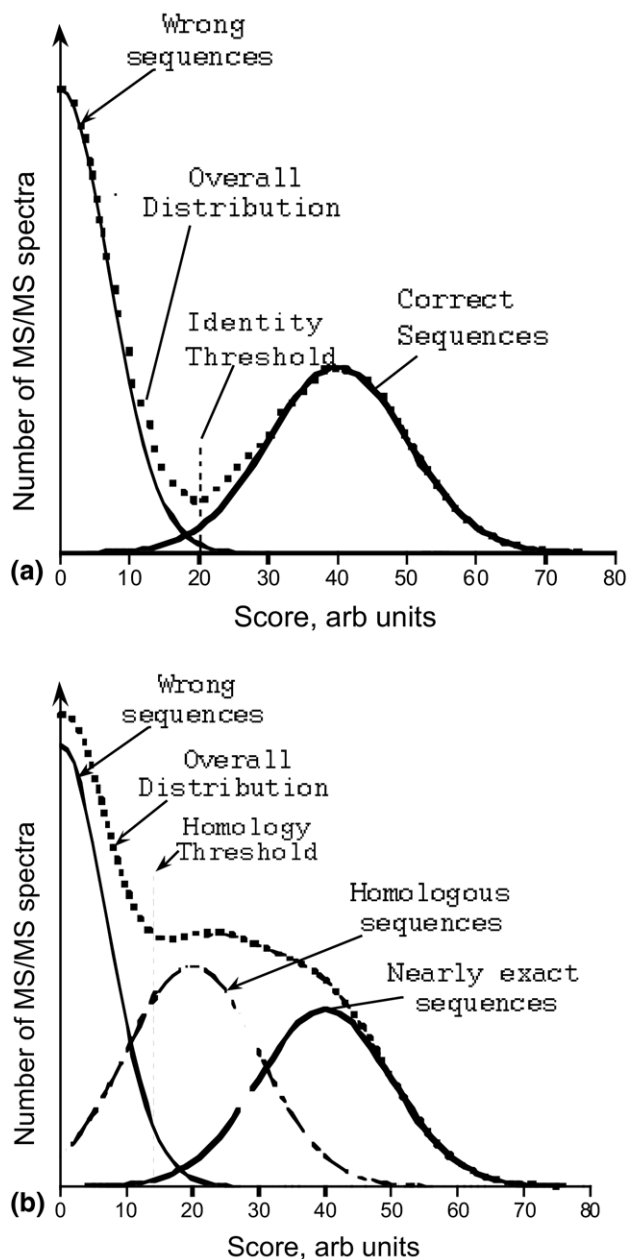
Another kind of MS/MS homology is when the alternative second (or larger order) sequence assignment to the same MS/MS spectrum receives a lower

score than the best assignment, but which is still above the threshold. In this situation there is a great temptation to believe the top-score assignment and ignore the lower-scoring alternatives. The inevitable consequence of this decision is the possibility that many assignments received the top score not because they are correct but simply because a better-scoring homolog was absent in the database. One way to test the validity of the received score is to compare it with an expected value derived from an independently assigned parameter, such as S-score [47], although even that method gives no guarantee. It is thus more logical to accept as “correct” all above-threshold sequence assignments, thus acknowledging the very simple truth that in the absence of complete MS/MS sequence information sufficient for full de novo sequencing the best of what a database search can establish is MS/MS homology and not sequence identity.

The presence of peptide MS/MS homologs makes it practically impossible to reach an error-free status in peptide identification from incomplete MS/MS information. On the idealized histogram of search-engine score values reflecting the goodness of the fit, there are two clearly separated distributions of the false positive and true positive IDs (Figure 7a) [48]. In such a “hit or



**Figure 6.** Distribution of the number of alternative peptide sequences (zero means unique sequence) assigned by Mascot with a top score above an automatically determined threshold value to each MS/MS dataset: (a) for search in the full NCBI database; (b) for search against the human database. Ten was the highest allowed number of sequences.



**Figure 7.** Theoretical shapes of the distributions of search-engine scores: (a) idealized “hit or miss” situation; (b) more realistic situation taking into account the existence of homologs MS/MS datasets.

miss” case, it is possible to set up a score threshold above which the assignments are considered reliable. In reality, however, there is an intermediate distribution (or rather, a distribution of distributions) of homolog hits that make it impossible to set up a meaningful border between the true hits from false positives (Figure 7b). At every threshold value, some number of accepted hits will be attributed to homolog peptides with sequences different from those of the assigned ones.

Another artifact causing multiple alternative assignments is the simultaneous selection for MS/MS of several peptides with similar  $m/z$  values, which can

occur when chromatographic resolution is insufficient and many peptides elute at the same time. We will not consider this problem in detail here, except for noting that more extensive sequence information usually helps to deal with it.

One possible solution to the MS/MS homology problem is relying not only on the absolute score of the best sequence assignment to an MS/MS dataset, but also on the score difference with the second-best assignment. This approach is implemented in Sequest, XTandem, new versions of Mascot (data on Figures 5 and 6 were obtained with an older Mascot version), and possibly other search engines. In this approach, if more than one sequence from a database can be assigned to an MS/MS spectrum, this assignment is not trusted. Although such an approach definitely increases confidence in sequence assignment, it cannot completely eliminate the homology problem and, depending on the size of the database, it can significantly limit the efficiency of a database search. Indeed, according to Figure 6, 20 to 64% of all MS/MS datasets with an above-threshold score are not to be trusted. If the database included all possible peptide sequences, few incomplete MS/MS datasets could be assigned with certainty because almost every incomplete dataset would be assignable to more than one peptide sequence.

In our opinion, a more radical solution of the MS/MS homology issue could be based not on disfavoring alternative assignments arising mostly because of incompleteness of sequence information in MS/MS datasets, but on elimination of this incompleteness altogether. Such elimination is equivalent to the ability of performing full de novo sequencing of every analyzed peptide. At the moment this is beyond the possibility of mass spectrometry in routine analysis, but in principle obtaining full sequence information on every analyzed peptide is a realistic goal. One should bear in mind, however, that even full de novo sequencing leaves unsolved the sequence homology problem that stems from the fact that many peptides, both tryptic and nontryptic, are found in more than one protein sequence.

### Role of Mass Accuracy

Reliability of de novo sequencing is not an easy issue by itself, and this topic is rarely discussed in the literature. MS/MS sequencing has scored rather poorly in a recent ABRF study where none of 106 participating laboratories could correctly sequence all five nontryptic peptides [49]. In our opinion, the reason for that poor performance is not the design of de novo algorithms, but the insufficient quality of mass spectrometry data, given that the truly high-performance de novo sequencing procedure with >95% reliability demands a combination of two complementary fragmentation techniques with ppm molecular mass accuracy and fragment mass accuracy of better than about 20 mDa [16]. The attempts to achieve the 95% reliability target using low-resolution CxD data

have so far been futile even when most sophisticated de novo sequencing methods have been recruited. For instance, the approach that uses several independently acquired Cx/D MS/MS spectra of the same or slightly modified sequence produced on average 50 hexapeptide sequence tags per spectrum, with only 82% probability of at least one tag to be correct [50]. Thus even sequence tag identification is not reliable enough if data are obtained with a low-resolution and single fragmentation technique. In contrast to that, a Reliable Sequence Tag (RST) derived from high-resolution Cx/D + Ex/D data is more than 98% correct [16, 47]. For more on the benefits of mass accuracy see Zubarev and Mann [51].

### How to Combine Cx/D and Ex/D

Combination of data from two complementary fragmentation techniques in one datafile can be done as described in Nielsen et al. [44]. An even better alternative would be to submit to database search two MS/MS spectra “as is.” Unfortunately, none of the existing search engines can, to the best of our knowledge, accept such information today. The informational content of side-chain losses, which are abundant in both techniques [18, 40, 52], is still underutilized by search engines, especially Ex/D losses [40].

## Conclusions

In this work we make a case for the combined use of Ex/D and Cx/D techniques in tandem mass spectrometry applied to shotgun proteomics. We argue that the only way to reliably identify peptide sequences is through de novo sequencing, which is achievable with desired >95% reliability only by using two complementary fragmentation techniques and high mass accuracy. Thus the answer to the question expressed in the title is: duet, and duet only.

## Acknowledgments

This work was supported by the Knut and Alice Wallenberg Foundation, the Swedish Research Council (grants 621-2004-4897 and 621-2003-4877 to RZ).

## References

- Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. Electron Capture Dissociation of Multiply Charged Protein Cations. A Non-ergodic Process. *J. Am. Chem. Soc.* **1998**, *120*, 3265–3266.
- Zubarev, R. A.; Fridriksson, E. K.; Horn, D. M.; Kelleher, N. L.; Kruger, N. A.; Carpenter, B. K.; McLafferty, F. W. Electron Capture Dissociation for Structural Characterization of Multiply Charged Protein Cations. *Anal. Chem.* **2000**, *72*, 563–573.
- Tsybin, Yu. O.; Håkansson, P.; Budnik, B. A.; Haselmann, K. F.; Kjeldsen, F.; Gorshkov, M.; Zubarev, R. A. Improved Low-Energy Injection Systems for High Rate Electron Capture Dissociation in Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2001**, *15*, 1849–1854.
- Haselmann, K. F.; Budnik, B. A.; Olsen, J. V.; Nielsen, M. L.; Reis, C. A.; Clausen, H.; Johnson, A. H.; Zubarev, R. A. Advantages of External Accumulation for Electron Capture Dissociation in Fourier Transform Mass Spectrometry. *Anal. Chem.* **2001**, *73*, 2998–3005.
- Palmblad, M.; Tsybin, Yu. O.; Ramstrom, M.; Bergquist, J.; Håkansson, P. Liquid Chromatography and Electron-Capture Dissociation in Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 988–992.
- Davidson, W.; Frego, L. Micro-High-Performance Liquid Chromatography/Fourier Transform Mass Spectrometry with Electron-Capture Dissociation for the Analysis of Protein Enzymatic Digests. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 993–998.
- Tsybin, Yu. O.; Wetterhall, M.; Markides, K. E.; Håkansson, P.; Bergquist, J. Capillary Electrophoresis and Electron Capture Dissociation Fourier Transform Ion Cyclotron Resonance Mass Spectrometry for Peptide Mixture and Protein Digest Analysis. *Eur. J. Mass Spectrom.* **2002**, *8*, 389–396.
- Baba, T.; Hashimoto, Y.; Hasegawa, H.; Hirabayashi, A.; Waki, I. Electron Capture Dissociation in a Radio Frequency Ion Trap. *Anal. Chem.* **2004**, *76*, 4263–4266.
- Silivra, O. A.; Kjeldsen, F.; Ivonin, I. A.; Zubarev, R. A. Electron Capture Dissociation in a Quadrupole Ion Trap: First Results. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 22–27.
- Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. Peptide and Protein Sequence Analysis by Electron Transfer Dissociation Mass Spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 9528–9533.
- Appella, E.; Anderson, C. W. New Prospects for Proteomics: Electron-Capture (ECD) and Electron-Transfer Dissociation (ETD) Fragmentation Techniques and Combined Fractional Diagonal Chromatography (COFRADIC). *FEBS J.* **2007**, *274*, 6255–6255.
- Khidekel, N.; Ficarro, S. B.; Clark, P. M.; Bryan, M. C.; Swaney, D. L.; Rexach, J. E.; Sun, Y. E.; Coon, J. J.; Peters, E. C.; Hsieh-Wilson, L. C. Probing the Dynamics of O-GlcNAc Glycosylation in the Brain Using Quantitative Proteomics. *Nat. Chem. Biol.* **2007**, *3*, 339–348.
- Taverna, S. D.; Ueberheide, B. M.; Liu, Y.; Tackett, A. J.; Diaz, R. L.; Shabanowitz, J.; Chait, B. T.; Hunt, D. F.; Allis, C. D. Long-Distance Combinatorial Linkage between Methylation and Acetylation on Histone H3N Termini. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 2086–2091.
- Molina, H.; Horn, D. M.; Tang, N.; Mathivanan, S.; Pandey, A. Global Proteomic Profiling of Phosphopeptides Using Electron Transfer Dissociation Tandem Mass Spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 2199–2204.
- Zubarev, R. A. Protein Primary Structure Using Orthogonal Fragmentation Techniques in Fourier Transform Mass Spectrometry. *Expert Rev. Proteomics* **2006**, *3*, 251–261.
- Savitski, M. M.; Nielsen, M. L.; Kjeldsen, F.; Zubarev, R. A. Proteomics-Grade De Novo Sequencing Approach. *J. Proteome Res.* **2005**, *4*, 2348–2354.
- Fälth, M.; Savitski, M. M.; Nielsen, M. L.; Kjeldsen, F.; Andren, P. E.; Zubarev, R. A. SwedCAD, a Database of Annotated High-Mass Accuracy MS/MS Spectra of Tryptic Peptides. *J. Proteome Res.* **2007**, *6*, 4063–4067.
- Huang, Y.; Triscari, J. M.; Tseng, G. C.; Pasa-Tolic, L.; Lipton, M. S.; Smith, R. D.; Wysocki, V. H. Statistical Characterization of the Charge State and Residue Dependence of Low-Energy CID Peptide Dissociation Patterns. *Anal. Chem.* **2005**, *77*, 5800–5813.
- Dongre, A. R.; Somogyi, A.; Wysocki, V. H.; Surface-Induced Dissociation: An Effective Tool to Probe Structure, Energetics and Fragmentation Mechanisms of Protonated Peptides. *J. Mass Spectrom.* **1996**, *31*, 339–350.
- Savitski, M. M.; Kjeldsen, F.; Nielsen, M. L.; Garbuzynskiy, S. O.; Galzitskaya, O. V.; Surin, A. K.; Zubarev, R. A. Backbone Carbonyl Basicities Relate Gas-Phase Fragmentation of Peptides and Protein Folding. *Angew. Chem. Int. Ed.* **2007**, *46*, 1481–1484.
- Paizs, B.; Suhai, S. Fragmentation Pathways of Protonated Peptides. *Mass Spectrom. Rev.* **2005**, *24*, 508–548.
- Budnik, B. A.; Nielsen, M. L.; Olsen, J. V.; Haselmann, K. F.; Hörth, P.; Haehnel, W.; Zubarev, R. A. Can Relative Cleavage Frequencies in Peptides Provide Additional Sequence Information? *Int. J. Mass Spectrom.* **2002**, *219*, 283–294.
- Savitski, M. M.; Kjeldsen, F.; Nielsen, M. L.; Zubarev, R. A. Complementary Sequence Preferences of Electron Capture Dissociation and Vibrational Excitation in Fragmentation of Polypeptide Poly-cations. *Angew. Chem. Int. Ed.* **2006**, *45*, 5301–5303.
- Tsybin, Y. O.; He, H.; ben Hamidane, H.; Emmett, M. R.; Hendrickson, C. L.; Tsybin, O. Y.; Marshall, A. G. Electron Capture/Transfer Dissociation Product Ion Abundances: Correlation with Amino Acid Hydrophobicity and Application in Peptide and Protein Structural Analysis. In *Proceedings of the 55th ASMS Conference on Mass Spectrometry and Allied Topics*, Indianapolis, Indiana, June 3–7, 2007.
- Sze, S. K.; Ge, Y.; Oh, H. B.; McLafferty, F. W. Top-Down Mass Spectrometry of a 29-kDa Protein for Characterization of any Posttranslational Modification to within One Residue. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 1774–1779.
- Ge, Y.; Lawhorn, B. G.; ElNaggar, M.; Strauss, E.; Park, J. H.; Begley, T. P.; McLafferty, F. W. Top Down Characterization of Larger Proteins (45 kDa) by Electron Capture Dissociation Mass Spectrometry. *J. Am. Chem. Soc.* **2002**, *124*, 672–678.
- Charlebois, J. P.; Patrie, S. M.; Kelleher, N. L. Electron Capture Dissociation and C-13, N-15 Depletion for Deuterium Localization in Intact Proteins after Solution-Phase Exchange. *Anal. Chem.* **2003**, *75*, 3263–3266.
- Kjeldsen, F.; Budnik, B. A.; Haselmann, K. F.; Jensen, F.; Zubarev, R. A. Dissociative Capture of Hot (3–13 eV) Electrons by Polypeptide Poly-cations: An Efficient Process Accompanied by Secondary Fragmentation. *Chem. Phys. Lett.* **2002**, *356*, 201–206.



29. Kjeldsen, F.; Sørensen, E.; Zubarev, R. A. Distinguishing of Ile/Leu Amino Acid Residues in the PP3 Protein by (Hot) Electron Capture Dissociation in Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Anal. Chem.* **2003**, *75*, 1267–1274.
30. Kelleher, N. L.; Zubarev, R. A.; Bush, K.; Furie, B.; Furie, B. C.; McLafferty, F. W.; Walsh, C. T. Localization of Labile Posttranslational Modifications by Electron Capture Dissociation: The Case of  $\gamma$ -Carboxyglutamic Acid. *Anal. Chem.* **1999**, *71*, 4250–4253.
31. Cooper, H. J.; Håkansson, K.; Marshall, A. G. The Role of Electron Capture Dissociation in Biomolecular Analysis. *Mass Spectrom. Rev.* **2005**, *24*, 201–222.
32. Zubarev, R. A. Electron Capture Dissociation Tandem Mass Spectrometry. *Curr. Opin. Biotechnol.* **2004**, *15*, 12–16.
33. Cooper, H. J.; Hudgins, R. R.; Håkansson, K.; Marshall, A. G. Characterization of Amino Acid Side Chain Losses in Electron Capture Dissociation. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 241–249.
34. Good, D. M.; Wirtala, M.; McAlister, G. C.; Coon, J. J. Performance Characteristics of Electron Transfer Dissociation Mass Spectrometry. *Mol. Cell. Proteomics* **2007**, *6*, 1942–1951.
35. Nielsen, M. L.; Budnik, B. A.; Haselmann, K. F.; Olsen, J. V.; Zubarev, R. A. Intramolecular Hydrogen Atom Transfer in Hydrogen-deficient Polypeptide Radical Cations. *Chem. Phys. Lett.* **2000**, *330*, 558–562.
36. Budnik, B. A.; Haselmann, K. F.; Zubarev, R. A. Electron Detachment Dissociation of Peptide Di-Anions: An Electron-Hole Recombination Phenomenon. *Chem. Phys. Lett.* **2001**, *342*, 299–302.
37. Chi, A.; Huttenhower, C.; Geer, L. Y.; Coon, J. C.; Syka, J. E. P.; Bai, D. L.; Shabanowitz, J.; Burke, D. J.; Troyanskaya, O. G.; Hunt, D. F. Analysis of Phosphorylation Sites on Proteins from *Saccharomyces cerevisiae* by Electron Transfer Dissociation (ETD) Mass Spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *14*, 2193–2198.
38. Savitski, M. M.; Kjeldsen, F.; Nielsen, M. L.; Zubarev, R. A. Hydrogen Rearrangement to and from Radical  $z$  Fragments in Electron Capture Dissociation of Peptides. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 113–120.
39. Lin, C.; Cournoyer, J. J.; O'Connor, P. B. Probing the Gas Phase Folding Kinetics of Peptide Ions by IR Activated DR-ECD. In *Proceedings of the 55th ASMS Conference on Mass Spectrometry and Allied Topics*, Indianapolis, Indiana, June 3–7, 2007.
40. Savitski, M. M.; Nielsen, M. L.; Zubarev, R. A. Side-Chain Losses in Electron Capture Dissociation Improve Peptide Identification. *Anal. Chem.* **2007**, *79*, 2296–2302.
41. Horn, D. M.; Zubarev, R. A.; McLafferty, F. W. Automated De Novo Sequencing of Proteins by Tandem High-Resolution Mass Spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 10313–10317.
42. [www.matrixscience.com/pdf/asms\\_tutorial\\_2005.pdf](http://www.matrixscience.com/pdf/asms_tutorial_2005.pdf).
43. Thorne, G. C.; Ballard, K. D.; Gaskell, S. J. Metastable Decomposition of Peptide  $[M + H]^+$  Ions via Rearrangement Involving Loss of the C-Terminal Amino Acid Residue. *J. Am. Soc. Mass Spectrom.* **1990**, *1*, 249–257.
44. Nielsen, M. L.; Savitski, M. M.; Zubarev, R. A. Improving Protein Identification Using Complementary Fragmentation Techniques in Fourier Transform Mass Spectrometry. *Mol. Cell. Proteomics* **2005**, *6*, 835–845.
45. Gignac, S. M.; Brauer, S.; Häne, B.; Quentmeier, H.; Drexler, H. G. Elimination of Mycoplasma from Infected Leukemia Cell Lines. *Leukemia* **1991**, *5*, 162–165.
46. Anderson, N. L.; Polanski, M.; Pieper, R.; Gatlin, T.; Tirumalai, R. S.; Conrads, T. P.; Veenstra, T. D.; Adkins, J. N.; Pounds, J. G.; Fagan, R.; Lobley, A. The Human Plasma Proteome: A Nonredundant List Developed by Combination of Four Separate Sources. *Mol. Cell. Proteomics* **2004**, *3*, 311–316.
47. Savitski, M. M.; Nielsen, M. L.; Zubarev, R. A. New Data Base-Independent, Sequence Tag-Based Scoring of Peptide MS/MS Data Validates Mowse Scores, Recovers below Threshold Data, Singles Out Modified Peptides, and Assesses the Quality of MS/MS Techniques. *Mol. Cell. Proteomics* **2005**, *8*, 1180–1188.
48. Nesvizhskii, A. I.; Keller, A.; Kolker, E.; Aebersold, R. A Statistical Model for Identifying Proteins by Tandem Mass Spectrometry. *Anal. Chem.* **2003**, *75*, 4646–4658.
49. Turck, C. W.; Falick, A. M.; Kowalak, J. A.; Lane, W. S.; Neubert, T. A.; Phinney, B. S.; Weintraub, S. T.; West, K. A. [http://www.abrf.org/ResearchGroups/Proteomics/EPsters/PRG05poster\\_revised\\_051212\\_stw.pdf](http://www.abrf.org/ResearchGroups/Proteomics/EPsters/PRG05poster_revised_051212_stw.pdf). In *Proceedings of the 53rd ASMS Conference on Mass Spectrometry and Allied Topics*, San Antonio, TX, 2005.
50. Bandeira, N.; Tsur, D.; Frank, A.; Pevzner, P. A. Protein Identification by Spectral Networks Analysis. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 6140–6145.
51. Zubarev, R. A.; Mann, M. On the Proper Use of Mass Accuracy in Proteomics. *Mol. Cell. Proteomics* **2007**, *6*, 377–381.
52. Savitski, M. M.; Kjeldsen, F.; Nielsen, M. L.; Zubarev, R. A. Relative Specificities of Water and Ammonia Losses from Backbone Fragments in Collision-Activated Dissociation. *yyj. Proteome Res.* **2007**, *6*, 2669–2673.